



THESIS

VALIDITY OF THE ONE-MILE WALK-TEST AS A PREDICTOR OF AEROBIC  
CAPACITY

Submitted by

ELIZABETH G. FONTENOT

Department of Physiology

In partial fulfillment of the requirements for the  
degree of Master of Science

Colorado State University  
Fort Collins, Colorado  
Spring 2001

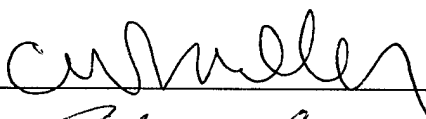
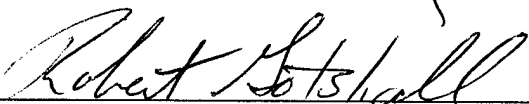
**THE VIEWS EXPRESSED IN THIS ARTICLE  
ARE THOSE OF THE AUTHOR AND DO NOT  
REFLECT THE OFFICIAL POLICY OR  
POSITION OF THE UNITED STATES,  
DEPARTMENT OF DEFENSE, OR THE U.S.  
GOVERNMENT**


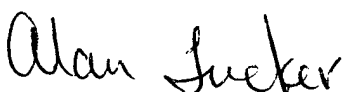
COLORADO STATE UNIVERSITY

March 1, 2001

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR  
SUPERVISION BY 2<sup>ND</sup> LT. ELIZABETH G. FONTENOT ENTITLED VALIDITY OF  
THE ONE-MILE WALK-TEST AS A PREDICTOR OF AEROBIC CAPACITY BE  
ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF  
MASTER OF SCIENCE.

Committee on Graduate Work

  
\_\_\_\_\_  
  
\_\_\_\_\_

\_\_\_\_\_  
  
\_\_\_\_\_  
Advisor  
  
\_\_\_\_\_  
Department Head/Director

## ABSTRACT

### VALIDITY OF THE ONE-MILE WALK-TEST AS A PREDICTOR OF AEROBIC CAPACITY

The United States Air Force is interested in finding easier and less expensive means for testing fitness of National Guard and Reserve members. A walk-test fits the criteria of being easy to accomplish, with minimal equipment, and with limited stress on the body. This study tested the hypothesis that the equation developed by Dolgener et al. (1994) for a 1-mile walk test is a reliable predictor of  $\text{VO}_2$  max in college age female military subjects. Thus, if the hypothesis is supported, the 1-mile walk test would meet the needs of the Air Force Reserve and National Guard for fitness testing of their personnel. The subjects resided at the United States Air Force Academy, which is at an altitude of 2,200 meters. The results of this study were that the actual  $\text{VO}_2$  max was  $41.83 \text{ ml kg}^{-1} \text{ min}^{-1}$  and the predicted  $\text{VO}_2$  max was  $57.66 \text{ ml kg}^{-1} \text{ min}^{-1}$ , indicating that the Dolgener equation significantly over-estimated  $\text{VO}_2$  max in this population ( $t\text{-stat} = 14.95$ ,  $t\text{-critical } 2\text{-tail} = 2.042$  and  $p < 0.001$ ). However, after consulting additional published equations from the Dolgener study, the best predictions were made by a gender specific equation that accounted for weight. This female specific equation which calculated  $\text{VO}_2$  max in  $\text{ml kg}^{-1} \text{ min}^{-1}$ , predicted a mean  $\text{VO}_2$  max value of  $38.057 \text{ ml kg}^{-1} \text{ min}^{-1}$ . In the Dolgener study, this equation was not presented as a reliable estimate for predicting  $\text{VO}_2$  max, but in this study it provided fairly accurate predictions on the specific population of the present study. Thus, this study demonstrates that in a young

military female population at altitude, a 1-mile walk test using the acceptable Dolgener equation developed on college-age males and females in 1994 does not accurately predict  $\text{VO}_2$  max in the gender specific population of the current study.

Elizabeth G. Fontenot  
Department of Physiology  
Colorado State University  
Fort Collins CO  
Spring 2001

## Table Of Contents

### CHAPTER

I. INTRODUCTION	1-3
II. LITERATURE REVIEW	4-29
III. METHODS	29-31
IV. RESULTS	32-40
V. DISCUSSION	41-45
VI. CONCLUSIONS AND RECOMMENDATIONS	46-47
VII. REFERENCES	48-51

## CHAPTER 1

### INTRODUCTION

Measuring the maximal oxygen consumption ( $\text{VO}_2 \text{ max}$ ) is the current standard that is used to estimate aerobic fitness in individuals. Direct measurement of  $\text{VO}_2 \text{ max}$  requires expensive equipment, high subject motivation, and is time-consuming (Kline et al. 1987). Due to these limitations, sub-maximal tests have been preferred by some to predict  $\text{VO}_2 \text{ max}$ . These sub-maximal tests include step, cycle ergometer, treadmill, walking, and running tests. Several researchers (Shoenfeld et al. 1981, Jackson et al. 1990) have attempted to predict  $\text{VO}_2 \text{ max}$  based on multiple regression equations using several factors, such as resting heart rate, body weight, height, and physical activity.

Physical fitness is important to the military and there is a continual interest to develop methods that can accurately determine the fitness of the troops. The U.S. Air Force (USAF) has employed a 1.5-mile run test for estimation of aerobic fitness for cadets at the U. S. Air Force Academy (USafa) and for cadets in the Reserve Officer Training Corps (ROTC). This method is not used with active duty personnel due to some safety concerns. An article in the Air Force Times (September 11, 1989) indicated that five individuals died from cardiac incidents related to the annual running aerobic test during a 12-month period. The military has thus been using the Air Force Cycle Ergometry Test (CET) for all its active duty personnel. However, there is a current impetus to implement the one-mile walk test to determine aerobic fitness for the Air Force Reserve and Air National Guard personnel.



The walk test is safer, less expensive, and less complicated for Reserve and National Guard personnel who are not always near military facilities where the CET test can be given. The walk test would then allow the personnel of these units to perform the annual sub-maximal test with little or no equipment. Kline et al. (1987) developed an equation for predicting  $\text{VO}_2$  max in (L/min) that is considered to be reliable for populations over the age of 30 yrs: given Kline's equation, Dolgener et al. (1994) derived a general equation that is considered to be reliable for prediction of  $\text{VO}_2$  max (l/min) in the population of 18 to 30-year-old individuals:

$$\text{VO}_2 \text{ max (l/min)} = 3.5959 + 0.0096 (\text{WT}) + 0.6566 (\text{SEX}) - 0.0096 (\text{TIME}) - 0.0080 (\text{HR})$$

WT = Weight (lbs)

SEX = 0: F, 1: M

TIME = time of walk (minutes)

HR = final heart rate (bpm)

A study by George et al. (1998) suggests that the equation of Dolgener et al. (1994) is an accurate predictor of  $\text{VO}_2$  max for college age subjects. A recent study by Chuba (2000) indicated that the equations of Dolgener et al. (1994) accurately predicted  $\text{VO}_2$  max for a predominately male population in this age group. Chuba (2000) suggested that a female population should be tested to determine if the equation accurately predicts  $\text{VO}_2$  max in women. Therefore, because the USAF has personnel in both the Reserve and National Guard that fit into the age group of 18-30 years and is a mixture of male and females, this research study was designed to test the reliability of this test for females in the age range of 18-30 yrs.

The purpose of this study was to compare the estimated  $\text{VO}_2$  max, using a sub-maximal 1-mile walk test, to a direct measure of  $\text{VO}_2$  max in a sample of female subjects

in the age group of 18-30 yrs. The hypothesis was that the field walk test would accurately predict the  $\text{VO}_2$  max of female military personnel of the age population 18-20 yrs old. Results were evaluated and compared using the regression equations previously developed (Dolgener et al., 1994; Kline et al., 1987).

## CHAPTER II

### LITERATURE REVIEW

There are many methods to measure physical fitness in an individual. Some methods include evaluating muscular endurance, flexibility, body composition, strength, and aerobic capacity. Aerobic capacity can be determined using a variety of different tests and can be used to predict physical fitness. There is some difficulty in measuring aerobic capacity, but according to Fitchett (1985) “the internationally accepted reference standard for cardiorespiratory fitness despite recent criticisms...is the maximum oxygen uptake ( $\text{VO}_2$  max).” Thus, when it is necessary to measure what Knapik (1989) describes as the “body’s ability to consume and utilize oxygen,” the  $\text{VO}_2$  max test is the route by which to achieve this measurement.

$\text{VO}_2$  max is defined to be the point where an individual’s oxygen uptake does not increase despite an increase in exercise intensity (Knapik 1989). This test is the gold standard test for measuring aerobic capacity, but it has some drawbacks. This test is expensive, time-consuming, and not practical for large subject numbers (Kline et al. 1987). In addition, this test is difficult for first-time subjects because the actions of breathing through a mask while running or cycling are unfamiliar. This unfamiliarity can cause these subjects to hyperventilate or request to stop the test before reaching their maximum potential. Due to these limitations, the  $\text{VO}_2$  max test is not always a practical method for measuring aerobic capacity of large groups.

The limitations associated with the  $\text{VO}_2$  max test have encouraged new methods to be developed, where an individual is exercised at a sub-maximal level and the  $\text{VO}_2$  max is estimated. The sub-maximal testing method is more practical for large subject groups and limited budgets. These tests are easier to administer and offer more flexibility for testing because the necessary equipment is usually less expensive and easier to transport. The USAF currently uses sub-maximal testing due to the large number of people that has to be tested every year. Examples of sub-maximal tests include step, walk, cycle ergometer, and running tests.

The USAF is considering the one-mile walk test for the Reserve and National Guard personnel. Sub-maximal testing offers the USAF an objective fitness test that is inexpensive, easy to administer, can be completed almost anywhere, and is relatively safe. In addition, this test would assist in helping to test personnel in the Reserve and National Guard who do not live near a permanent base. These individuals are currently forced to make a yearly trip to a base for the current physical fitness test using the cycle ergometer. This test would then save the personnel from the inconvenience of missing work and traveling long distances to a USAF base which has the ergometer equipment.

The one-mile walk test was determined by Kline et al. (1987) to be a valid predictor of aerobic capacity for individuals over the age of 30. This test was then determined to be invalid for individuals less than 30 yrs of age by Dolgener et al. (1994). Kline and colleagues (1987) derived correlations between predictions for their regression equations to actual  $\text{VO}_2$  max measurements to be high ( $r = 0.85-0.88$ , SEE of  $4.5-5.0 \text{ ml kg}^{-1} \text{ min}^{-1}$ ). When a cross-validation was done in the Kline study with their equations it was found that there were "no significant differences between the observed and estimated

VO<sub>2</sub> max values for any of the cross-validation analyses.” (Kline et al. 1987). Thus, the conclusions of this study pointed to the walk test as being a useful tool to predict VO<sub>2</sub> max for all individuals over the age of 30 yrs. When Dolgener and colleagues (1994) used the original Kline regression equations on college age individuals, they found that these equations significantly ( $t = 9.95$ ,  $p < 0.001$ ) over-predicted the VO<sub>2</sub> max of the subjects.

Due to the discrepancy between age groups, the USAF would like a validation of the one-mile walk test as a predictor of VO<sub>2</sub> max for individuals between the ages of 18-30 because of the large number of personnel in the Reserve and National Guard that fall into this age group. To understand the purpose of this experiment, it will be necessary to first present some background information defining the maximal oxygen uptake test, the current information on the one-mile walk test, and the information on the sub-max cycle ergometer test currently employed by the USAF.

## THE MAXIMAL OXYGEN CONSUMPTION TEST

### *Physiology*

Mitchell et al. (1957) described the maximal oxygen uptake test as being “when one subjects a normal individual to progressively increasing workloads, allowing sufficient time for recovery between each increment of work, a linear relationship between workload and oxygen intake is found.” Upon continuing to increase the workload, the maximal oxygen intake per unit time will not increase with the workload. It is at this point that a maximal test is complete. This test illustrates that there is positive relationship between oxygen intake and workload. In addition, this study suggests that maximal oxygen uptake relates to cardiac output and the ability of the tissues to extract

oxygen from the blood. This study demonstrated that as a subject changes from a resting state to a heavy workload, which is going to produce the  $\text{VO}_2$  max, that the subject's oxygen intake increased 9.5 times, cardiac output increased 4.3 times, and arterial-venous oxygen ( $\text{AVO}_2$ ) difference increased by 2.2 times. Thus, the Mitchell study concluded that there is a dependent relationship between maximal oxygen intake and both cardiac output and  $\text{AVO}_2$  difference. Mitchell et al. (1957) were especially interested in the widening of the  $\text{AVO}_2$  difference by 2.2 times because it allowed the oxygen intake to exceed 3 liters/min. This measurement suggested to the researchers that if it was not for an organism's ability "to widen its AV oxygen difference, cardiac output would have to increase nearly 10 times to supply 3,200 ml of oxygen per minute to the tissues."

#### *Limitations of the Maximal Aerobic Test*

The maximal oxygen uptake  $\text{VO}_2$  max test is the "Gold Standard" by which aerobic capacity is measured, but it has its drawbacks. The level of training of the individual, the type of test that is performed, and the mechanics of the testing method are just a few factors that can limit the usefulness of the  $\text{VO}_2$  max test. Other limitations are that the test is time-consuming, expensive, and hard to use on large groups of people. Cardiac patients face the undue risk of increased stress from the physical exertion that is expected in order to achieve maximal oxygen uptake (Rowell 1964). Harrison et al. (1980) stated that for an individual who is not trained in cycling, "factors unrelated to the delivery and utilization of  $\text{O}_2$  can influence  $\text{VO}_2$  during cycling, and that, therefore, the bicycle ergometer may not be suitable for the direct determination at  $\text{VO}_2$  max."

In addition to problems with the maximal tests, Harrison et al. (1980) suggest that the  $\text{VO}_2$  max test may not be a sufficient measure to base calibrations of indirect testing

or sub-maximal testing. Harrison et al. (1980) also suggest that it is difficult to measure or predict  $\text{VO}_2$  max using one test. Therefore, more than one  $\text{VO}_2$  max test should be conducted on subjects and the mean  $\text{VO}_2$  max calculated. In the study (Harrison et al., 1980), five subjects completed several  $\text{VO}_2$  max tests and the 95% tolerance interval for a single measurement of  $\text{VO}_2$  max was  $\pm 7.8\%$ . This information led the authors to suggest that the range of variability around a single  $\text{VO}_2$  max measurement is similar to the average increase in  $\text{VO}_2$  max that can be achieved by physical training. Therefore, it may be difficult to attribute improved  $\text{VO}_2$  max values to training. Harrison et al. (1980) went on to estimate that physical training would only improve  $\text{VO}_2$  max by 7 to 33 percent.

#### *Specificity of Testing*

Test specificity can influence the measurement of an individual's  $\text{VO}_2$  max. Athletes specifically trained in a certain mode of exercise can attain higher  $\text{VO}_2$  max using that specific exercise (Fernhall et al., 1990). Fernhall et al. (1990) measured  $\text{VO}_2$  max in runners and cyclists using both a treadmill and a cycle ergometer. It was shown that both the cyclists and runners had higher  $\text{VO}_2$  max results on the treadmill test (mean of  $59.6 \text{ ml/kg}^{-1}/\text{min}^{-1}$ ) versus the cycle ergometer (mean of  $52.6 \text{ ml/kg}^{-1}/\text{min}^{-1}$ ). The cyclists showed a 7.4% higher  $\text{VO}_2$  max on the treadmill compared to the cycle ergometer, and the runners showed a 16% higher  $\text{VO}_2$  max on the treadmill compared to the cycle ergometer. This difference between the runners'  $\text{VO}_2$  max on the treadmill and the cycle ergometer was significant, but the difference between  $\text{VO}_2$  max on the treadmill and the cycle ergometer for the cyclists were not significant. However, the cyclist  $\text{VO}_2$  max values on the cycle ergometer were significantly higher than for the runners on the same test. Therefore, the treadmill test elicited higher  $\text{VO}_2$  max values for both runners

and cyclists, although for the cyclists, the difference in  $\text{VO}_2$  max between the two testing methods was not significant. The fact that the cyclists did not perform better on the cycle ergometer test did not agree with previous research, but Fernhall et al. (1990) accounted for their testing difference by stating that other studies used elite cyclists and their study did not. McArdle et al. (1996) stated that in non-specifically trained athletes the treadmill registered higher  $\text{VO}_2$  max values than the cycle ergometer test.

Harrison et al. (1980) recommend using the treadmill test for “calibrating” indirect methods against direct methods. This theory is based on a study performed with 10 male subjects who performed both the cycle ergometer  $\text{VO}_2$  max test and a running uphill treadmill test for  $\text{VO}_2$  max at 10 km/hr and 12 km/hr. No significant difference were observed between the two treadmill tests performed at different speeds, but the  $\text{VO}_2$  max on the cycle ergometer was 20% lower than on the treadmill ( $p < 0.001$ ). The literature indicates that for most athletes the treadmill test is the most reliable method to test  $\text{VO}_2$  max, unless the athlete is specifically trained in cycling.

Despite its limitations, the  $\text{VO}_2$  max test is still the most reliable method to determine aerobic capacity. The guidelines to determine when a subject is nearing his/her  $\text{VO}_2$  max include a plateau of the  $\text{O}_2$  uptake versus workload relationship, approaching the subject's maximum heart rate ( $220 - \text{age}$ ), fatigue exhibited by the subject, and the knowledge that the subject is primarily burning carbohydrates for fuel as seen with an RQ value  $\geq 1.0$  (McArdle et al., 1996).

The determination of the primary energy source of a subject is by using the respiratory exchange ratio (RER). This measurement is made by determining alveolar levels of  $\text{CO}_2$  and  $\text{O}_2$  and determining the ratio of  $\text{CO}_2$  produced to  $\text{O}_2$  consumed. It is



reflective of the respiratory quotient (RQ), which is a measure of the  $\text{CO}_2/\text{O}_2$  ratio produced at the cellular level. Both of these measurements are useful tools in research of this nature. Because of the different chemical compositions of lipids, proteins, and carbohydrates, different amounts of oxygen are used to oxidize these compounds to water and carbon dioxide. The  $\text{CO}_2/\text{O}_2$  ratio differ for each substrate being oxidized being  $\approx 1.0$  for carbohydrates,  $\approx 0.7$  for lipids, and  $\approx 0.8$  for proteins. (McArdle et al 1996).

Therefore, the RER provides an indication of what fuel the subject is primarily oxidizing during exercise. It is considered that when a subject is approaching  $\text{VO}_2$  max, the RER exceeds 1.0 due to the oxidation of carbohydrates as the primary source of fuel. The use of the RQ value as a measure of physical fitness was introduced by Issekutz et al. (1962). It was determined by these researchers that “the relative increase in  $\text{CO}_2$  output compared to the  $\text{O}_2$  uptake was the result of accumulation of lactic acid with a concomitant decrease of the body bicarbonate pool.” The study used 32 untrained subjects and had them perform a 5-minute test on a cycle ergometer where the subjects pedaled at a constant frequency of 50 rev/min and a constant workload. The test was repeated with increasing workloads until the subject achieved a maximal  $\text{O}_2$  uptake. Through the use of Douglas bags at the 3.5-minute mark for 45-60 seconds, the researchers measured  $\text{CO}_2$  output and  $\text{O}_2$  intake. This allowed the researchers to track the RQ values. The results indicated that when the RQ value was 1.15 the subjects had reached their maximal  $\text{O}_2$  uptake. This study then illustrated the use of RQ as a useful tool to approximate maximal  $\text{O}_2$  uptake.

#### SUBMAXIMAL TESTS TO PREDICT AEROBIC CAPACITY

In the 1920's, investigators at the Harvard Fatigue Laboratory attempted to develop a sub-maximal test to classify individual fitness levels. The test required

individuals to drag a weighted sled 300 yards. Upon completion of dragging the sled, the individuals had their recovery heart rate measured. This first sub-maximal test was called the "the stone boat test." This test later evolved into the Harvard Step Test (Kline et al. 1987). The step test, in addition to bike tests, walk tests, and run tests, are just a few of the methods that have been developed to measure aerobic capacity without a maximal effort. Each test has advantages and limitations in its ability to calculate aerobic capacity. Through comparison of the different tests, some conclusions have been reached. The best tests for predicting aerobic capacity appear to be the run and the step tests; however, these tests require motivation and can pose some health risks for less fit individuals. A less valid test that poses less risk to the health of individuals is the bike test; whereas, the walk test appears to be more accurate and less risky than the bike tests. The major criteria to determine which test to use depends on equipment availability and the number of subjects to be evaluated.

Most sub-maximal tests depend on the ideas presented by Astrand and Ryhming (Rowell et al., 1964). They developed a nomogram to predict  $\text{VO}_2$  max based on the theory that  $\text{VO}_2$  max and heart rate are linearly related over a wide range of values. Thus, the max  $\text{VO}_2$  can be predicted by extrapolation of the slope of sub-maximal  $\text{VO}_2$  versus heart rate to an assumed maximal heart rate (Rowell et al. 1964). Astrand and Ryhming based the nomogram on a study in which subjects, between the ages of 20 and 30 years of age and in good health, performed at a set workload while heart rate was monitored. The group of healthy male subjects ( $n=50$ ) had a heart rate of 128 bpm after 6 minutes at 50% of the  $\text{VO}_2$  max. For a corresponding group of female subjects ( $n=62$ ), the heart rate was 138 bpm. At 70% of the  $\text{VO}_2$  max the subjects achieved heart rates of 154 bpm and 164

bpm, respectively, for males and females. From this information, Astrand and Ryhming developed a nomogram to predict aerobic capacity. The claim was made that the nomogram could predict aerobic capacity when tested on the cycle ergometer and step test to within 6% in 2/3 of the subjects tested (Astrand and Ryhming, 1954). The test worked best, according to the authors, when the workload was high enough to get a steady heart rate between 125 and 170 bpm because they suggest that, within these limits, almost linear increases in metabolism and heart rate occur. The results may be skewed due to hypoxia, hot climates, and dehydration that will result in a higher heart rate at a given intensity (Astrand and Ryhming, 1954).

The Astrand and Ryhming nomogram was adjusted for age by Teraslinna et al. (1966), criticized the Astrand-Ryhming nomogram for suggesting that the maximum cardiac output can be reached. Their study found that  $\text{VO}_2$  could still continue to rise beyond what the nomogram shows as max cardiac output. Thus this study illustrated that the nomogram does not illustrate where maximum cardiac output is reached because cardiac output can continue to increase despite the nomogram's results. This rise in  $\text{VO}_2$  will still continue to increase because, for a given work rate,  $\text{O}_2$  intake approaches the asymptote more slowly than does heart rate. Hence, the nomogram would underestimate the  $\text{VO}_2$  max since it is dependent only on maximal heart rate. Fitchett (1985) indicated that the linear relationship between heart rate and  $\text{VO}_2$  max breaks down at higher workloads because the heart rate often reaches a maximum before  $\text{O}_2$  uptake is maximal.

Despite the criticisms associated with the original prediction methods for  $\text{VO}_2$  max, predictions are necessary because of the limitations of the  $\text{VO}_2$  max test. Several methods have been developed for predicting  $\text{VO}_2$ , but they are mainly based on two

assumptions. First, that heart rate is linearly related to oxygen uptake over a wide range of workloads, and second, that all subjects within a population age group are capable of reaching similar maximal heart rates (Fitchett 1985). Because indirect measurements are highly variable, these methods should be judged against the standard  $\text{VO}_2$  max test to determine validity (Harrison et al. 1980). Variability can occur among individuals of varying physical ability. It has been shown that fewer errors occur between predicted and actual values of  $\text{VO}_2$  max in athletes versus sedentary people, and the margin of difference decreases with physical training (Rowell 1964).

### *Step Test*

The step test has existed for a long time and through the years has evolved into various protocols, such as the Harvard Step test and the Queens College Step test. These tests basically require an individual to step on and off a bench of a determined height for a set period of time. The height of the bench and the length of time are set by the protocol. In a comparison study of the Queens College Step test to other sub-maximal tests, it was determined that the Queens College Step test produced a significantly lower estimate of  $\text{VO}_2$  max than either the 1.5-mile run or the 1-mile walk tests (Zwiren et al. 1991). In another comparison study by Harrison et al. (1980), it was suggested that the 2-km run and the step test were the best estimates of  $\text{VO}_2$  max when compared to the bike and walk tests. It was also suggested in this study that increasing the step height and lowering the step frequency may increase the accuracy of the  $\text{VO}_2$  max prediction. Harrison et al. (1980) attributed this to the greater comfort and ease with which the test could be preformed. This procedure is referred to as the modified step test because the height was modified from the optimal 40-cm step for subjects with leg lengths of 75-85-

cm to 45-cm step for subjects with longer leg lengths. It was also noted that both the run and step test required 60% of the subject's aerobic power which could increase the medical risk to certain subjects.

In conclusion, Fitchett (1985) suggested that the step test is a valid measure of aerobic capacity for population studies but is susceptible to considerable error of prediction when used on an individual basis. The literature on the step test indicates that it is as reliable a test of aerobic capacity as any of the other sub-maximal tests, but in its modified form may be one of the best predictors in those individuals who have little risk of cardiovascular disease.

### *Run Tests*

Cooper (1968) developed the first running test, a 12-minute run, that was used to predict  $\text{VO}_2$  max. This test was determined to be a "highly reliable and valid indicator of maximum oxygen intake" (Doolittle and Bigbee 1968). Both Doolittle and Bigbee (1980) and Cooper (1968) found a correlation coefficient of 0.90 between the predicted 12-minute run  $\text{VO}_2$  max and measured  $\text{VO}_2$  max in 9<sup>th</sup> grade boys. Maksud et al. (1971) attempted to apply this run to young males (ages 13-14, approximately 8<sup>th</sup>-9<sup>th</sup> grades) and determined a 0.65 correlation coefficient and advised it be used with caution.

Researchers reasoned that since exercise prescription to improve aerobic capacity requires at least 15 minutes of aerobic activity, most of the run tests are too short and hence have an anaerobic component. In pursuit of developing a test for strictly aerobic capacity, Jackson et al. (1990) analyzed the validity of a 3-mile run as a test for aerobic capacity in college males. The data from this study indicated that the 3-mile run only presents a moderate correlation ( $r=0.58$ ) to  $\text{VO}_2$  peak. Since shorter runs have a

high concurrent validity, there seems to be no reason to replace the shorter run tests with a 3-mile run. However, one point that does favor the 3-mile run test is that the construct validity (discrimination between groups of different aerobic capacities) of the 3-mile run is not present in the shorter runs. Thus, this makes the 3-mile run acceptable for those subjects who have the fitness level to endure the test. The U. S. Army uses a 2-mile run to evaluate aerobic fitness of their troops because of the research that supports the high correlation ( $r = -0.76$  to  $-0.91$ ) between  $\text{VO}_2$  max and 2-mile run times (Knapik 1989).

Comparison of the run test to other sub-maximal test indicates that the run tests are the most accurate predictors of aerobic capacity, but are often too strenuous to apply to the general population. Harrison et al. (1980) reported comparisons of the step, walk, bike, and run tests and suggested that, for subjects free of cardiac abnormalities, the run test provides the best estimate of aerobic capacity. As for individuals with health problems or handicaps this same study indicated that the walking test was safer but less reliable compared to the run test. Harrison et al. (1980) indicated that the 2-km run predictions of  $\text{VO}_2$  max were superior to those of the 3, 5, and 7-km walks. However, this run test did require a high motivation on the part of the subjects. Thus, Harrison et al. (1980) agreed with Cooper (1968) that the run test would be best for military personnel since motivation would be considered equally as important as fitness in this unique population.

In comparing the run test to cycling tests, Glassford et al. (1965) indicated that the muscle mass used in running is much greater than that used in cycling which contributes to the excessive fatigue of specific muscles and less oxygen utilized during cycling tests.

This may provide an explanation for the superiority of the running tests when compared to the bike tests.

### *Bike Tests*

The bike test is beneficial because it does not require the motivation associated with the run tests, nor is it as stressful. It is a test that can be employed for use with all types of subjects. The only criticism is that it has not been shown to be as reliable as run or walk tests in predicting  $\text{VO}_2$  max. It has been shown in comparison experiments that both the walk and run tests, and even the step tests, are more reliable in predicting  $\text{VO}_2$  max than the bike test. Zwiren et al. (1991) indicated that the bike test, which uses the Astrand and Ryhming nomogram to estimate  $\text{VO}_2$  max from heart rate measured at a fixed sub-maximal power output, was a less reliable method than the run or walk tests. Zwiren et al. (1991) went on to illustrate that the bike test significantly underestimates  $\text{VO}_2$  max. The explanation for this underestimate was attributed to the limitation of the asymptotic nature of the heart rate- $\text{VO}_2$  curve at heavier workloads, that leads to a larger than expected increase in  $\text{VO}_2$  per unit increase in heart rate. Thus, the  $\text{VO}_2$  would increase by a larger amount than the increase in heart rate would indicate. Zwiren et al. (1991), however, indicated that an overestimation of  $\text{VO}_2$  from the indirect bike test could be explained by the magnitude of individual differences in the oxygen cost of cycle ergometer exercise. Also, all the subjects of this study were females, and Astrand (1960) illustrated that females compared to males attained a lower oxygen uptake by 400-500 ml  $\text{min}^{-1}$  at a given workload.

Despite the unreliability of the bike test, it is still a valid predictor of  $\text{VO}_2$  max for population studies; however, Fitchett (1983) suggested that the bike test might be subject

to error for individual assessments. Glassford et al. (1965) reported that the direct treadmill tests and the indirect Astrand and Ryhming bicycle test had higher mean values than the direct Astrand and Ryhming bicycle test. It was explained that the subject pool that was used to create the indirect Astrand and Ryhming nomogram were elite athletes who could elicit greater cardiovascular responses. It was shown by Fitchett (1983) that the indirect Astrand and Ryhming bike test is a good predictor of aerobic capacity because no significant difference was observed reported between the maximal treadmill test and the indirect Astrand and Ryhming bike test. Patton et al. (1982) noted that the indirect Astrand and Ryhming bike test correlated significantly with the direct cycle ergometer test and the treadmill test, but only for the male subjects. This could be attributed again to the difference between males and females in oxygen uptake per given workload, but still does not explain the difference with regard to the treadmill test. This suggests that the Astrand and Ryhming indirect bike test may be gender specific, and would not be a good measure for populations such as in the military, which require a gender-neutral test. Myles et al. (1982) suggested that a different indirect bike test of maximal aerobic power be used on healthy males and compared to a direct treadmill measurement of  $VO_2$ . This indirect test, the Indirect Maximal Aerobic Power test (IMAP), differs from the Astrand and Ryhming test as it requires a maximal effort and correlates more reliably with  $VO_2$  max. The IMAP test may be useful in categorizing individuals into fitness levels, which is the goal of the military.

Fernall et al. (1990) indicated that since less muscle mass is used in cycling, it is possible that heart rate is influenced by lesser blood flow to the contracting muscles. Cyclists exhibit a consistently lower heart rate during sub-maximal work compared with



runners. Fernhall et al. (1990) conducted a study to determine the effect of specificity of exercise training on the results of treadmill and cycle ergometry tests. It was found that the level of training is a factor in sub-maximal testing. The  $\text{VO}_2$  max was not significantly different between runners and cyclists during the treadmill and cycle ergometer max tests, but runners had significantly lower  $\text{VO}_2$  max on the cycle ergometer test than the treadmill test, and cyclists had slightly higher  $\text{VO}_2$  max on the cycle ergometer test than on the treadmill. In the sub-maximal testing phase, the runners had significantly higher  $\text{VO}_2$  max on the sub-maximal running test than the cyclists. It was found that there was no significant difference in  $\text{VO}_2$  on the cycling sub-maximal test between runners and cyclists. This contrasts with the maximal test results which illustrated that trained cyclists performed better at cycling than did the runners. This appears to apply only to the maximal test. Therefore, a cycling sub-maximal test does not appear to be as good a predictor as a run test for those not specifically trained in cycling. In addition, it appears to make less of a difference as to which sub-maximal test is used for those trained in cycling.

Sub-maximal tests which use multiple regression equations, such as the run and walk tests, appear to be better predictors of  $\text{VO}_2$  max than those which rely on heart rate at a constant power load, like the bike test. It is possible that the bike test may be more accurate if it were to use more variables for its  $\text{VO}_2$  max prediction equations. Jessup et al. (1974) suggest that using multiple regression equations, as opposed to simple regression equations, may improve results in populations that are homogeneous. When they tested a homogeneous population with the Astrand-Ryhming nomogram in

comparison to multiple regression equations, the correlation coefficients increased from 0.64 to 0.81.

### *Walk Tests*

Walk tests are becoming popular in patients who suffer from cardiac or lung disease because they do not pose as much risk as with other sub-maximal tests. The 6-minute walk has proven to be a valid test of aerobic capacity for patients with chronic heart failure, as these patients perform activities at work-loads close to their maximum ability (Roul et al. 1998). It was suggested by Roul et al. (1998) that the walk test is good as a first-line screening test for these patients. The 6-minute walk was a good predictor when compared to  $\text{VO}_2$  max for those patients with end-stage lung disease, with the distance ambulated being the strongest correlation. When variables of age and weight were added, the correlation increased from 0.74 to 0.83 (Cahalin et al. 1995).

Enright et al. (1998) developed regression equations to predict the total distance healthy individuals could walk in six minutes. The results of this study suggested that the equations should be corrected for age, height, weight, and gender, since all of these factors have an impact on the distance walked. Oja et al. (1991) compared 2-km, 1.5-km and 1-km walk tests and found that the 2-km test was the most accurate for determining the cardiorespiratory fitness of healthy adults. This 2-km test included in its regression equation the variables of heart rate, age, and anthropometric values, because the extra variables explained 66-76% of the variance of predicted  $\text{VO}_2$  max, with a standard error of estimation of 9-15% of the mean. However, this study did note that the equation did not work well for a very active and fit population of men. Harrison et al. (1980) suggested that the walk test provided a poorer correlation to  $\text{VO}_2$  max than the run tests,

but that the slower the walk and the higher the time, the more accurate the prediction of  $\text{VO}_2$  max. Harrison et al. (1980) went on to indicate that as subjects walked quickly, they felt unnatural and this unnatural movement and feeling contributed to heart rate variability and poor correlations with treadmill max tests. However, Zwiren et al. (1991) suggest that because the walk test uses several variables to predict  $\text{VO}_2$  max, the subject could walk at several speeds without affecting the accuracy of the prediction. In addition, Zwiren et al (1991) suggest that the prediction of  $\text{VO}_2$  max for females aged 30 to 39 yrs. was most accurately predicted by the 1.5-mile run and the 1-mile walk. It appears from the literature that a 1 or 1.5 mile distance is the best for the walk test.

## U. S. AIR FORCE BIKE TEST

### *History of the Test*

The military has a long history of physical fitness testing because of the nature of the profession. Muscular endurance, muscular strength, and aerobic capacity are all parts of physical fitness, and these aspects are all tested in physical fitness tests. The U. S. Army fitness test, which claims to promote combat readiness, consists of three items: push-ups, sit-ups, and a 2-mile run. The basis for this test stems from research that demonstrates a high correlation between  $\text{VO}_2$  max and 2-mile run times (Knapik 1989). The USAF started aerobic fitness testing in the 1970s and the initial test consisted of a 1.5 mile run. This was based on the research of Cooper et al. (1968) on the 12-minute run. After 1992, the USAF switched from the run to a modified version of the Astrand-Ryhming cycle ergometer test. The reason for this change was a concern for safety.

In 1991, the USAF decided to decrease the allowable time of completion of the 1.5 mile run. Sharp (1991) indicated that this time change would cause a problem for

most USAF personnel. Sharp had discovered that only 40% of USAF personnel exercised regularly, and predicted that 50% would pass the old 1.5-mile standards, but that only 33% would pass the new standards. In addition, the USAF had some concerns regarding cardiac safety with the 1.5-mile run test. A review of all fatal cardiac events between January 1981 and December 1982 indicated that there were three deaths related to the aerobic exercise testing. All were individuals over the age of 35 and none were regular exercisers. This review by the USAF indicated that an average of 100 active-duty individuals die each year from coronary artery disease at times other than annual aerobic testing (Data from Biometrics Division, The Air Force Surgeon General's Office, Sharp 1991). A report in the Air Force Times, September 11, 1989, indicated that during one 12-month period five individuals died from incidents related to annual aerobic testing; only one was a regular exerciser (Sharp 1991). Sharp also reported from these data that the risk of fatal myocardial infarction is between 1 and 5 per 470,000 tests.

Sharp went on to devise a screening medical questionnaire that could reduce the risk of cardiac problems in individuals participating in the fitness test. This questionnaire considered "low risk" individuals to be under the age of 30 and who exercised regularly, whereas the "high risk" individuals were those over the age of 30, hypertensive, smoking, non-exercising males with angina. Based on this questionnaire, Sharp concluded that "if all those needing medical evaluation were considered equally spread about the Air Force, each facility would need to evaluate 1.5 subjects per work day—certainly not a large work load." Sharp suggest that 7 of the 8 individuals who died due to the fitness test would not have participated directly in the annual testing without a prior supervised fitness program had they participated in his study. Therefore, Sharp suggested a need for

more emphasis on physical fitness programs in addition to proper screening techniques before annual exercise testing.

In 1992, The USAF switched to a modified version of the Astrand-Ryhming Cycle Ergometry Test and called it the Air Force Cycle Ergometry Test (CET) in order to use a safer test that was still a valid predictor of  $\text{VO}_2$  max. The initial test consisted of 6-10 minutes of pedaling on a stationary cycle ergometer at the rate of 50 rpm while monitoring heart rate. The workload resistance is adjusted according to the heart rate. A computer program calculates the estimated  $\text{VO}_2$  max using the variables of age, gender, height, weight, heart rate change and workload (Chin 1996). The current test, implemented in 1999, is longer in duration with more moderate workload changes.

#### *Problems with the CET Test*

As the previously cited research suggests, the bike test may present less risk to the individual. However, the test's validity and accuracy are in question because the bike test as a predictor of aerobic capacity, utilizing the testing method of changes in heart-rate with a constant power load, is known to be a less accurate test. The original USAF CET appeared to underestimate  $\text{VO}_2$  max. Hartung et al. (1993) showed that the bike test underestimates  $\text{VO}_2$  max by 20%, but that the test is valid because of its consistent and equal underestimation of both unfit and fit subjects. Williford et al. (1994) found that the bike test underestimated  $\text{VO}_2$  max by about 17%. This equates to 68% of the officers being tested to have their predicted  $\text{VO}_2$  put them into the wrong fitness category. In addition, there will be 5% of the officer population that will be put into at least two categories below their actual fitness group. This is not an accurate enough test for the USAF.

Another study conducted by Lockwood et al. (1997) indicated that the USAF CET validity was “highly questionable” since the bike test underestimated  $\text{VO}_2$  max by 15%. This study was the only study to compare the USAF CET protocol with other common cycle ergometry protocols. In a comparison of the USAF test and another similar cycle test protocol (Progressive Cycle Ergometry Test, PROG protocol) they found the CET test to be less reliable than the PROG test. The CET test estimated  $\text{VO}_2$  max ranging from overestimates of 1.8% to underestimates of 17.3%. A single USAF CET test was shown in this study to be unreliable with an intraclass correlation coefficient for reliability of a single test of 0.26 (Lockwood et al. 1997). In a comparison of the sensitivity of the USAF CET and the PROG test, the PROG test was most sensitive (82.2%) and the USAF test was least sensitive (50.3%). Sensitivity for the PROG test was better for subjects with  $\text{VO}_2$  max values over  $35 \text{ ml kg}^{-1} \text{ min}^{-1}$ , whereas the USAF test was more sensitive for less fit individuals, or those with  $\text{VO}_2$  max values less than  $35 \text{ ml kg}^{-1} \text{ min}^{-1}$ . Based on this study, it was suggested that the USAF CET test was highly questionable and thus may not be a good method to predict fitness in military personnel.

#### ROCKPORT WALK TEST

##### *Development of the Test*

Kline et al. (1987) developed a 1-mile sub-maximal walk test which estimates  $\text{VO}_2$  max from the variables: heart rate, age, gender, body weight, and time that it takes to complete the 1-mile walk. The study used a total of 343 subjects, 165 males and 178 females, between the ages of 30 and 69 yrs. All the subjects were healthy and were not taking medication which could affect the heart rate responses. The subjects performed two 1-mile walk tests on separate days. The instructions were to walk as fast as possible,

without running, for 1-mile. The two walks had to be within 30 seconds of each other or the subject had to walk again until this criterion was met. The subjects were then divided into two groups. One group was the variable group that was used to develop the regression equation and the second group was assembled to cross-validate the regression equation. The walk times at the end of each quarter mile were recorded, in addition, heart rates were recorded every minute of the walk. The mean of the last two 1-min HRs at the end of each one-quarter mile segment for each walk was used in the equation. The following regression equation to estimate  $\text{VO}_2$  max was then compared to an actual treadmill  $\text{VO}_2$  max test:

$$\text{VO}_2 \text{ max} = 6.9652 + (0.0092 \cdot \text{WT}) - (0.0257 \cdot \text{AGE}) + (0.5955 \cdot \text{SEX}) - (0.2240 \cdot \text{T1}) - (0.0115 \cdot \text{HR1-4})$$

WT = weight in lbs

HR = mean heart rate at the end of each quarter mile in bpm

AGE = years

SEX = 0:f, 1:m

T1 = time to complete walk (minutes)

The group that was used to cross validate this equation produced an r-value of 0.092 and an SEE of  $0.335 \text{ l min}^{-1}$ . No difference between observed  $\text{VO}_2$  ( $2.66 \pm 0.94 \text{ l min}^{-1}$ ) and estimated ( $2.67 \pm 0.87 \text{ l min}^{-1}$ )  $\text{VO}_2$  was found. Kline et al. (1987) cross-validated this equation according to age by decade and found correlations that ranged from 0.89 – 0.93 and SEE's from  $\pm 0.278 - 0.356 \text{ l min}^{-1}$ . The investigators also cross-validated the equation for male and females separately and found SEE's of  $0.277$  and  $0.249 \text{ l min}^{-1}$ , respectively. No differences between observed and estimated mean  $\text{VO}_2$  max values were found for males or females using the sex-specific equations. The investigators also cross-validated this equation in  $\text{ml kg}^{-1} \text{ min}^{-1}$  and found similar correlations and SEE's, with no significant differences between observed and estimated  $\text{VO}_2$  max. With this

information, the investigators concluded that this 1-mile Rockport Walk Test is a valid and reliable sub-maximal test to predict  $\text{VO}_2$  max in a population of 30 to 69 years of age. In addition, it was found that a single walk test was sufficient to predict  $\text{VO}_2$  max.

#### *The Test and the College Age Population*

Studies on college age subjects (18-29 years of age) indicated that the original Kline equation fails with this population. Dolgener et al. (1994) tested the original Kline equation on 274 healthy Caucasian students (129 males, 145 females). The study divided the students into two groups. The first group of 196 subjects (100 females, 96 males) was randomly selected to serve as the validation sample. The remaining 78 subjects were used to cross-validate any new prediction equations that developed from the validation sample. Using the validation sample, it was found that the Kline equation overestimated  $\text{VO}_2$  max by 16% to 18% in the males and by 22% to 23% in the females. Therefore, a regression equation was developed by Dolgener et al. (1994) that utilized the same variables as the Kline equation except that age (determined to not have much effect on the outcome) was dropped. In addition, another modification was to use the final heart rate in accordance with research by Wilkie et al. (1987) which indicated that recovery heart rate is acceptable for use in this testing. Dolgener developed a multitude of equations to predict  $\text{VO}_2$  max in populations of males and females under the age of 30. The Dolgener equation accepted for both males and females that calculates  $\text{VO}_2$  max in L/min is as follows:

$$\text{VO}_2 \text{ max} = 3.5959 + (0.0096 * \text{WT}) + (0.6566 * \text{SEX}) - (0.0996 * \text{T1}) - (0.0080 * \text{HR})$$

SEX = 0:f, 1:m

WT = lbs

T1 = Time to complete walk (minutes)

HR = Final heart rate (bpm)



This equation, when cross-validated with the second group, resulted in a correlation of  $r = 0.84$  and  $SEE = 0.397 \text{ lmin}^{-1}$ . Dolgener concluded that the subject population in his study might have been less fit than those of the Kline study and this is the reason why the Kline study overestimates  $\text{VO}_2 \text{ max}$ . Kline et al. (1987) did suggest that the Rockport Walk Test would overestimate the  $\text{VO}_2 \text{ max}$  of low and medium fit individuals by 6% and 3.4%, respectively, and underestimate the  $\text{VO}_2 \text{ max}$  values of the high fit individuals by up to 8%. Kline et al. also suggest caution when using the equation to measure the extremes of a population.

A study by George et al (1998) attempted to modify the Rockport Walk Test for the college age population (20-29 years of age). In this study, they determined if  $\text{VO}_2 \text{ max}$  could be predicted from a  $\frac{1}{4}$ -mile walk at a self-selected brisk pace rather than a 1-mile walk at a maximal pace. The subjects for this study were between the ages of 18 and 29 years. The protocol for the test was to use a modified Rockport Walk Test, in which the subjects walked at a self-selected brisk (less than maximal) pace for 1-mile on a track that had  $\frac{1}{4}$ -mile markings. The subjects also performed a maximal test on the treadmill. The study compared the  $\frac{1}{4}$ -mile split results to the 1-mile walk end-time using both the original Kline equation and the modified Dolgener equation. It was discovered that the Kline equation over-predicted  $\text{VO}_2 \text{ max}$  by 8 to 20% for the 1-mile walk test in this subject population. Correlations were found to range from  $r = 0.64$  to  $0.84$ . The percentage of the predicted  $\text{VO}_2 \text{ max}$  falling within  $4.5 \text{ ml kg}^{-1} \text{ min}^{-1}$  of the actual  $\text{VO}_2 \text{ max}$  measurement using the  $\frac{1}{4}$  mile data ranged from 18 to 26%, whereas it was 35 to 64% for the 1-mile data. In contrast, the Dolgener equation slightly underestimated the  $\text{VO}_2 \text{ max}$  of this subject population. Correlations from this equation ranged from  $r = 0.65$

to 0.85. Using the percentage of predicted  $\text{VO}_2$  max within  $4.5 \text{ ml kg}^{-1} \text{ min}^{-1}$  of actual  $\text{VO}_2$  max ranged from 80-82% for the  $\frac{1}{4}$  mile data and 75-82% for the 1-mile data. Thus, George et al. (1998) concluded that the Kline equation over-predicts  $\text{VO}_2$  max (8-20%) in college-age individuals, confirming the conclusion of Dolgener et al. (1996). This study also suggests that the subjects may walk at any pace as long as they maintain a constant walk pace and achieve an exercise intensity within the linear portion of the heart rate- $\text{VO}_2$  relationship—which they suggest according to Golding et al. (1989) begins at 110 bpm.

It appears that the Rockport Walk Test is a reliable measure of aerobic fitness. In some populations, such as those aged 18 to 29 years, the original Kline equation needs revision so that it may be more accurate. Due to the inaccurate predictions of  $\text{VO}_2$  max from the Rockport Walk Test, Dolgener concluded that this test should not be used in a college population. Subjects should be able to walk at a self-selected pace as long as heart rate is higher than 110 bpm. It is possible that a shorter distance than 1-mile could be employed, but more research is needed in this area. In addition, researchers should be aware of fitness extremes in the populations which could skew the results when using this test.

## SUMMARY

The  $\text{VO}_2$  max test is the most accurate test for evaluating aerobic capacity for individuals. These maximal tests are time-consuming, costly, and not feasible for large populations. Thus, several sub-maximal tests were developed to account for the problems associated with the maximal tests, yet still predict  $\text{VO}_2$  max reliably. the multiple regression equations that are used to predict  $\text{VO}_2$  max by a sub-maximal test

appear to provide the most accurate predictions. The running tests prove to be the most accurate at predicting  $\text{VO}_2$  max. This holds true especially in populations where motivation is a factor. This is a main reasons the military has implemented running tests to evaluate the fitness of their personnel. However, running tests do present drawbacks that must be addressed. There is the stress placed on the participant and this stress can make theses test impractical for cardiac patients or for persons unaccustomed to strenuous exercise. Unfortunately, this was an issue for USAF members in the recent past and now the USAF has employed the bike test for safety reasons to replace the 1.5-mile run test for its troops. A cheaper and a more practical test is needed for Reserve and National Guard bases. This has increased the USAF interest in a 1-mile walking test. If the USAF plans to use this modified 1-mile walk test for Reserve and National Guard personnel, there is a need for more information on the test's effectiveness in estimating  $\text{VO}_2$  max in the female population under 30 years of age.

## CHAPTER III

### METHODS

#### PARTICIPANTS

Forty female subjects were selected on a volunteer basis from the United States Air Force Academy (USAFA) cadet and officer population. All subjects were between the ages of 18 and 30 years and, prior to data collection, each person signed an informed consent form approved by the U. S. Air Force Academy Institutional Review Board Committee and the Colorado State University Human Research Committee. Pregnant women (self reported) were not accepted. Subjects were selected on the basis that they should meet the USAF physical fitness standards.

#### EQUIPMENT

##### *Walk Test*

The sub-maximal walk test was performed on an indoor course around the cadet gymnasium. Heart rates were monitored using Polar Electro Vantage NV heart rate transmitters and watch receivers. Subjects were weighed on a Detecto digital scale and body fat measurements were taken with calipers using three sites (triceps, suprailiac, and thigh). Body fat composition was recorded for each subjects using tables that estimate body fat percentages based on the age of the subject and the sum of the three skinfold measurements. The data were then evaluated using the Microsoft SPSS statistics program and Excel spreadsheet. This program calculated the predicted statistics and

plotted the aerobic capacity using the equations that were developed by Dolgener et al. (1994).

#### *VO<sub>2</sub> max Test*

VO<sub>2</sub> max was measured using a Sensormedics VMAX system and motorized treadmill. Polar Vantage NV heart rate monitors recorded heart rate and the subjects VO<sub>2</sub> max was determined to be the highest value attained following completion of the test.

#### PROCEDURES

Subjects reported to the Human Performance Laboratory at the USAFA Cadet Gymnasium on two separate occasions. The laboratory is located at an approximate elevation of 2200 m (7258 ft.) above sea level. The subjects performed the VO<sub>2</sub> max test on a separate day than the walk test in order to eliminate fatigue effects. The two tests were separated by at least 24 hours. The order of the testing was determined randomly for each subject.

#### *Walk Test*

Body fat measurements were taken using the calipers and age and weight were self-reported before the walk. This was not used in the equation but was to provide more information on the subject population. Following the procedures of Kline et al. (1987), the subjects were asked to walk 1-mile as fast as possible without running. Heart rate was measured throughout the walk and the recovery heart rate (the heart rate upon crossing the 1-mile marker) was used in the equation. Wilkie et al (1987) indicated that this modification to the Kline equation was more acceptable and more practical. The course was described to the subjects and the researcher counted laps and recorded time and heart rate at the end of the walk.

### *VO<sub>2</sub> max Test*

The VO<sub>2</sub> max test was conducted by having the subjects walk and then run on a motorized treadmill with a steady rate but increasing grade. The criteria for a subject to achieve VO<sub>2</sub> max on the treadmill were the same as used by Kline et al. (1987): 1) Leveling off of oxygen consumption despite an increase in work; 2) Respiratory exchange ratio (RER) > 1.1; and 3) Heart rate no less than 15 beats below age-predicted maximal heart rate (220-age). In addition, the subject could stop the test at any point due to fatigue. The testing procedure followed the standard test used by the USAFA Human Performance Laboratory personnel. This protocol consisted of a 2-minute walk at 2.0 mph with a grade of 0%, then an increase in speed until 6.0 mph is reached with the grade remaining at 0%. The treadmill speed is maintained at this grade for two minutes. Then, at constant speed of 6.0 mph, the grade is increased by 2% every minute up to 10% grade. After 10% grade is reached, the grade is increased by 1% every minute until the subject is no longer able to continue the test.

### *Statistics*

Standard descriptive statistics were used to compare the sub-maximal walk test to the VO<sub>2</sub> max test data. Using Microsoft Excel, a Pearson Correlation was determined. A Standard paired t-test and a regression was calculated also on Excel.

## CHAPTER IV

### RESULTS

Female subjects (n=31), ranging in age from 18-29 years, participated in this study (Table 1). All subjects met the American College of Sports Medicine screening requirements for exercise testing. All subjects were associated with the military; 1 ROTC cadet, 2 Air Force Captains, 1 USAFA faculty member, and 27 USAFA cadets. Table 2 compares the descriptive data of the subjects in the current study and those previously studied by Chuba (2000) and Dolgener et al. (1994). It can be seen that the study population was very similar to the subjects in the Dolgener study. The women in the present study weighed less than the Dolgener population but all other variables were similar to those of the Dolgener group. The Chuba (2000) study population, which was predominately male, did have a higher body weight and  $\text{VO}_2$  max (ml/kg/min). Despite these differences, the females in the present study were comparable to subjects in the Chuba (2000) study for other variables. Table 3 presents the descriptive data of the cross-validation group used in the Dolgener study and the subjects in the Chuba (2000) and the present studies. The Dolgener mixed-gender validation and cross-validation groups had a mean  $\text{VO}_2$  max that was very similar to the females in the current study, while Chuba's predominately male cadet population registered a higher mean  $\text{VO}_2$  max (Figures 1 and 2). Figure 1 is a comparison of mean  $\text{VO}_2$  max values of the validation and cross-validation groups in the Dolgener study, illustrating that  $\text{VO}_2$  max (determined from the validation group) reliably predicted  $\text{VO}_2$  max in the cross-validation group. Figure 2

compares the mean VO<sub>2</sub> max and predicted VO<sub>2</sub> max of the subjects in the current study with those in Chuba and Dolgener studies, using the Dolgener equation. Again, this illustrates that the women in the present study were similar to those in the Dolgener study, from which the equation was developed. This also illustrates that VO<sub>2</sub> max was predicted quite accurately in the Chuba (2000) study. Figure 3 is a comparison of the three studies with the subjects' actual VO<sub>2</sub> max and the subjects predicted VO<sub>2</sub> max from the Dolgener equation. This figure illustrates that the Dolgener equation accurately predicted the VO<sub>2</sub> max of the subjects in both their cross-validation group and in the Chuba (2000) male cadet study. However, the Dolgener equation over-estimated VO<sub>2</sub> max by 37.8% in the present female population. All of the current subjects met the requirement of having an exercise heart rate of at least 110 bpm. However, the correlation between heart rate and walk time was determined in order to assess motivation of the subjects. As seen in Figure 4, the negative correlation suggests that some subjects underachieved, as it appears that they did not put their best effort into the walk.

TABLE 1 Subject Characteristics (mean  $\pm$  SD)

	<u>Average</u>	<u>SD</u>	<u>Max</u>	<u>Min</u>
Number of Subjects	31			
Age (yrs)	21.0	2.7	29.0	18.0
Weight (kg)	63.29	8.27	91.16	47.62
Body Fat (%)	17.89	3.92	22.70	10.00
VO <sub>2</sub> max (ml/kg/min)	41.83	5.65	57.60	32.00
Predicted VO <sub>2</sub> (ml/kg/min)	57.66	4.95	72.84	43.27
HR Walk (bpm)	148.13	20.48	190.00	114.00
Time Walk (min)	13.27	1.14	15.55	10.95



TABLE 2. Descriptive data from the three studies (mean  $\pm$  SD)

	<u>Dolgener</u>	<u>Chuba</u>	<u>Present*</u>
Number of Subjects	196	30	31
Age (yrs)	19.4 $\pm$ 2.74	21.4 $\pm$ 2.12	21.0 $\pm$ 2.70
Weight (kg)	68.1 $\pm$ 11.71	70.17 $\pm$ 8.84	63.29 $\pm$ 8.27
VO <sub>2</sub> max (ml/kg/min)	41.2 $\pm$ 8.09	45.04 $\pm$ 6.60	41.83 $\pm$ 5.65
HR Walk (bpm)	152 $\pm$ 21	136.8 $\pm$ 24.3	148.13 $\pm$ 20.48
Time Walk (min)	13.39 $\pm$ 1.08	12.96 $\pm$ 1.09	13.27 $\pm$ 1.14

\* All subjects are Females

TABLE 3. Descriptive Data from the three studies (mean  $\pm$  SD) with the cross-validation group from the Dolgener study

	<u>Dolgener</u>	<u>Chuba</u>	<u>Present*</u>
Number of Subjects	78	30	31
Age (yrs)	19.2 $\pm$ 2.30	21.4 $\pm$ 2.12	21.0 $\pm$ 2.70
Weight (kg)	68.2 $\pm$ 9.50	70.17 $\pm$ 8.84	63.29 $\pm$ 8.27
VO <sub>2</sub> max (ml/kg/min)	40.3 $\pm$ 6.49	45.04 $\pm$ 6.60	41.83 $\pm$ 5.65
HR Walk (bpm)	154 $\pm$ 20.2	136.8 $\pm$ 24.3	148.13 $\pm$ 20.48
Time Walk (min)	13.36 $\pm$ 1.22	12.96 $\pm$ 1.09	13.27 $\pm$ 1.14

\* All subjects are Females

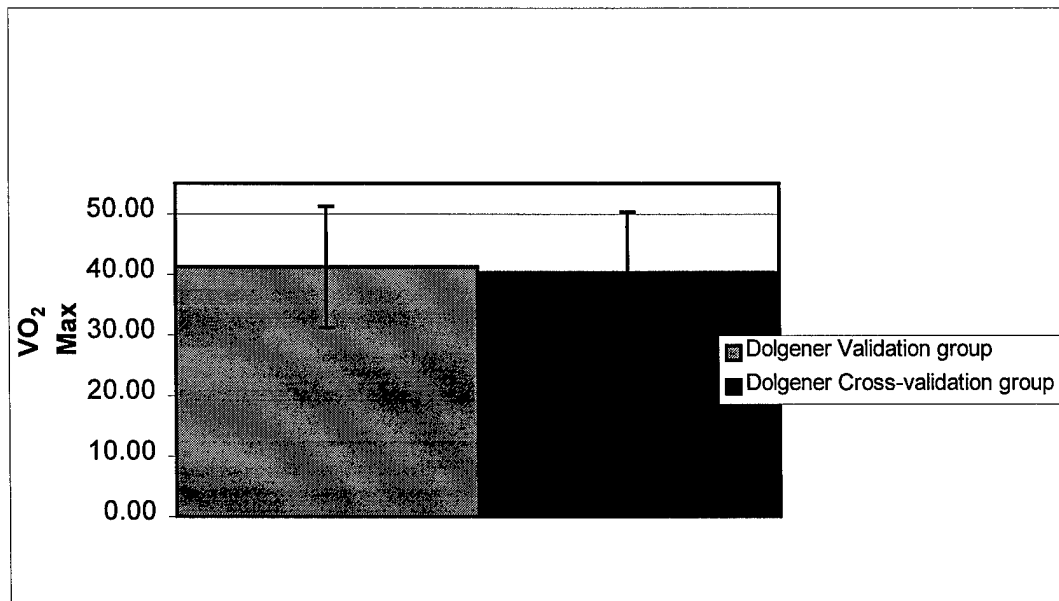


Figure 1: Dolgener Study VO<sub>2</sub> Max Comparisons of Validation vs. Cross-validation groups (ml kg<sup>-1</sup> min<sup>-1</sup>)

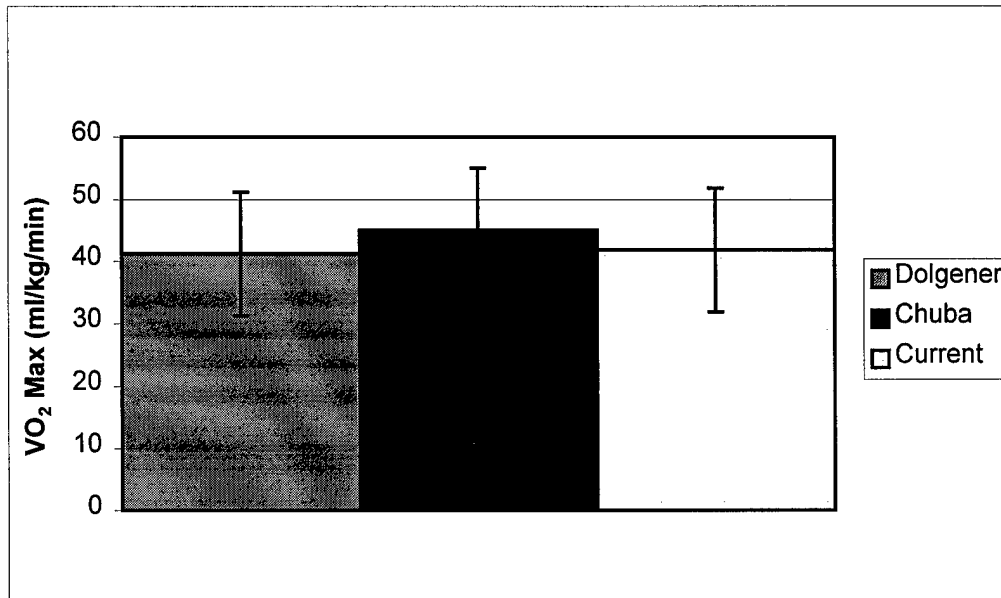


Figure 2: Average VO<sub>2</sub> Max of Subjects in Three Studies (ml kg<sup>-1</sup> min<sup>-1</sup>)  
 \*Dolgener Average VO<sub>2</sub> max is from the Validation group

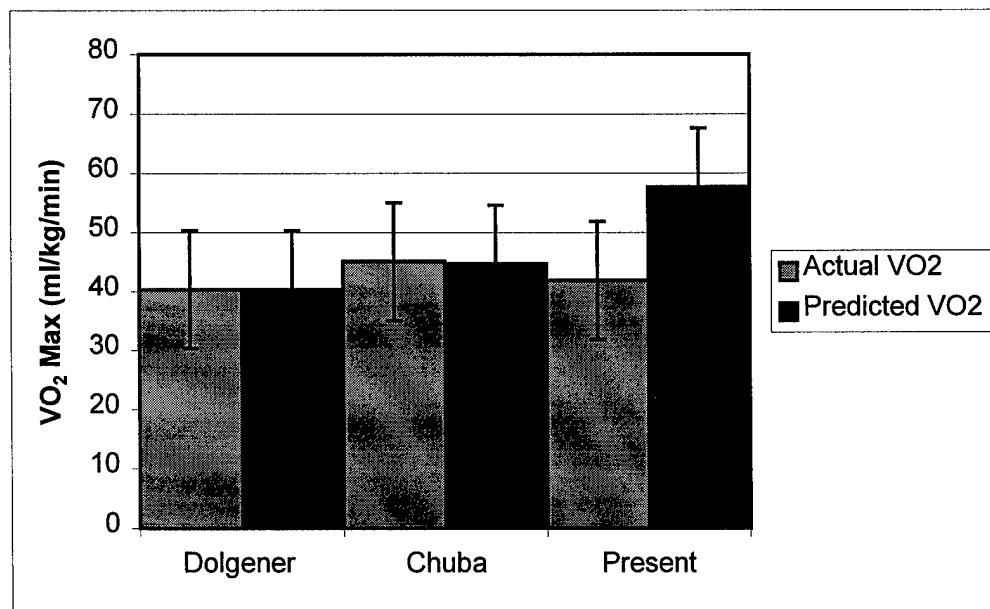


Figure 3: Average Actual VO<sub>2</sub> Max and Average Predicted VO<sub>2</sub> Max of Subjects in Three Studies (ml kg<sup>-1</sup> min<sup>-1</sup>)  
 \*Dolgener Actual and Predicted VO<sub>2</sub> max values from the Cross-validation group in the study.

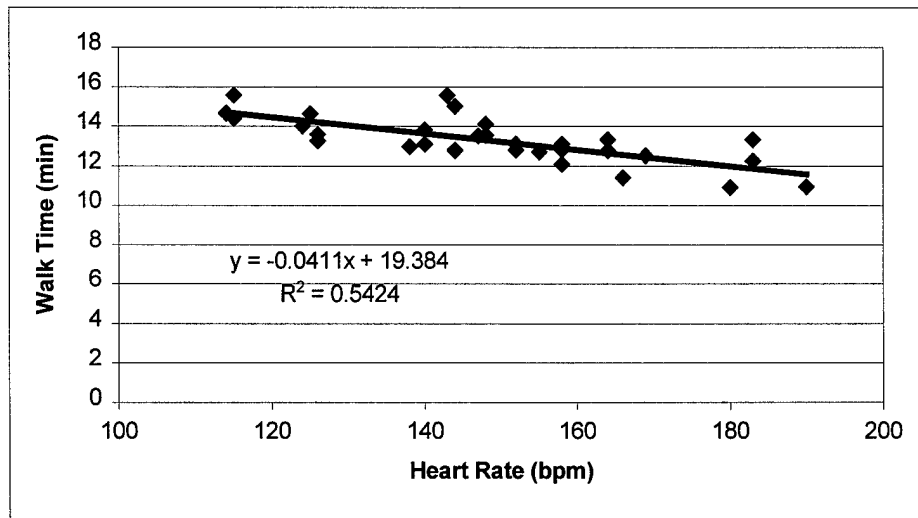


Figure 4: Scatter Plot of Heart Rate vs. Walk Time

## STATISTICAL INTREPRETATION

Using the Dolgener et al. (1994) equations, the predicted values were compared using paired samples t-tests. The paired samples t-test for the general Dolgener equation for both males and females calculated in  $L \min^{-1}$ , and then adjusted to  $ml \text{ kg}^{-1} \min^{-1}$ , indicated that with a high t-stat of 14.9, a t-value of 2.04, and a p-value  $< 0.001$  that the Dolgener equation significantly overestimated  $VO_2 \text{ max}$  in the present female population. The Pearson correlation between the predicted values of the Dolgener et al. (1994) equation and actual  $VO_2 \text{ max}$  indicated an  $R = 0.3881$ . Figure 5 is a scatter plot of the current study's actual  $VO_2 \text{ max}$  vs. the predicted  $VO_2 \text{ max}$  using the Dolgener equation compared to a line of index. This scatter plot illustrates that the Dolgener equation in this study consistently over-estimated predicted  $VO_2 \text{ max}$ .

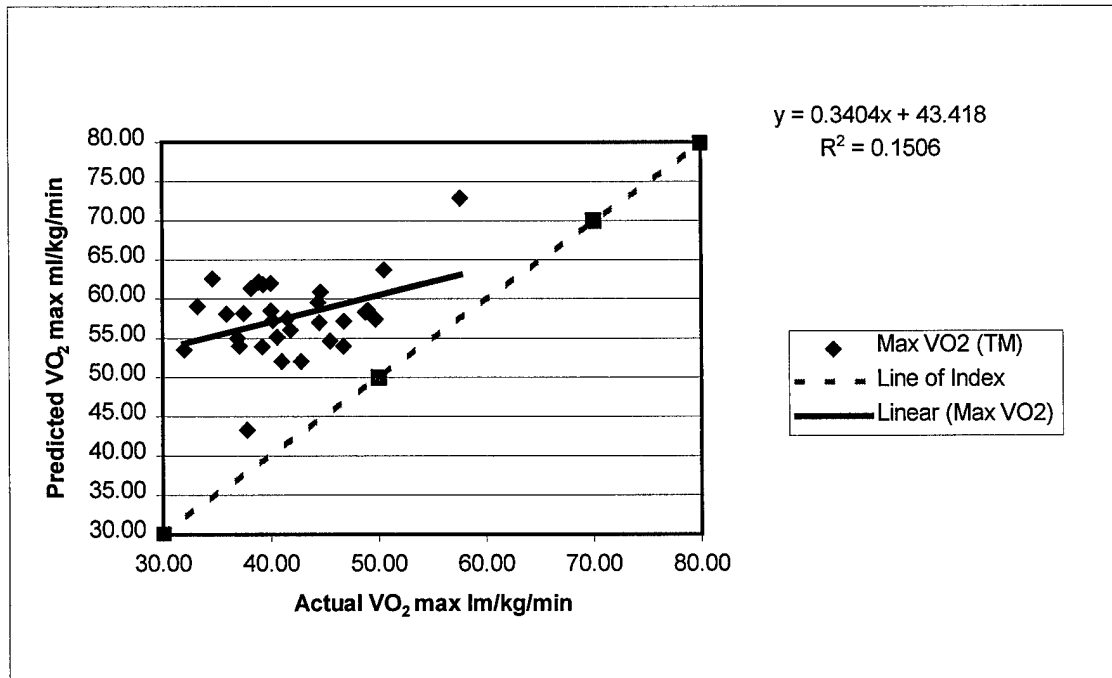


Figure 5: Actual VO<sub>2</sub> max vs. Predicted VO<sub>2</sub> Max (ml kg<sup>-1</sup> min<sup>-1</sup>)

As illustrated in Figure 5, every subject's VO<sub>2</sub> max was over-predicted by this equation. Since Dolgener et al. (1994) developed several equations in their study, but did not recommend them for predictions of VO<sub>2</sub> max, an attempt was made to evaluate these additional equations in the current population. The first equation evaluated was the female-specific equation for VO<sub>2</sub> max calculated in L min<sup>-1</sup>. This equation is:

$$\text{VO}_2 \text{ max} = 2.8328 + (0.0098 \cdot \text{WT}) - (0.0670 \cdot \text{T1}) - (0.0063 \cdot \text{HR})$$

WT = lbs

T1 = Time to complete walk (minutes)

HR = Final heart rate (bpm)

This equation's mean predicted VO<sub>2</sub> max for the current female population was 2.374 L min<sup>-1</sup>, which converts to 37.74 ml kg<sup>-1</sup> min<sup>-1</sup>. Both of these values are lower than the actual values of 2.63 L min<sup>-1</sup> and 41.83 ml kg<sup>-1</sup> min<sup>-1</sup>, respectively. A scatter plot of this

equation's predicted results versus the actual results (converted to  $\text{ml kg}^{-1} \text{min}^{-1}$ ) show the underestimation of  $\text{VO}_2 \text{ max}$  (Figure 6).

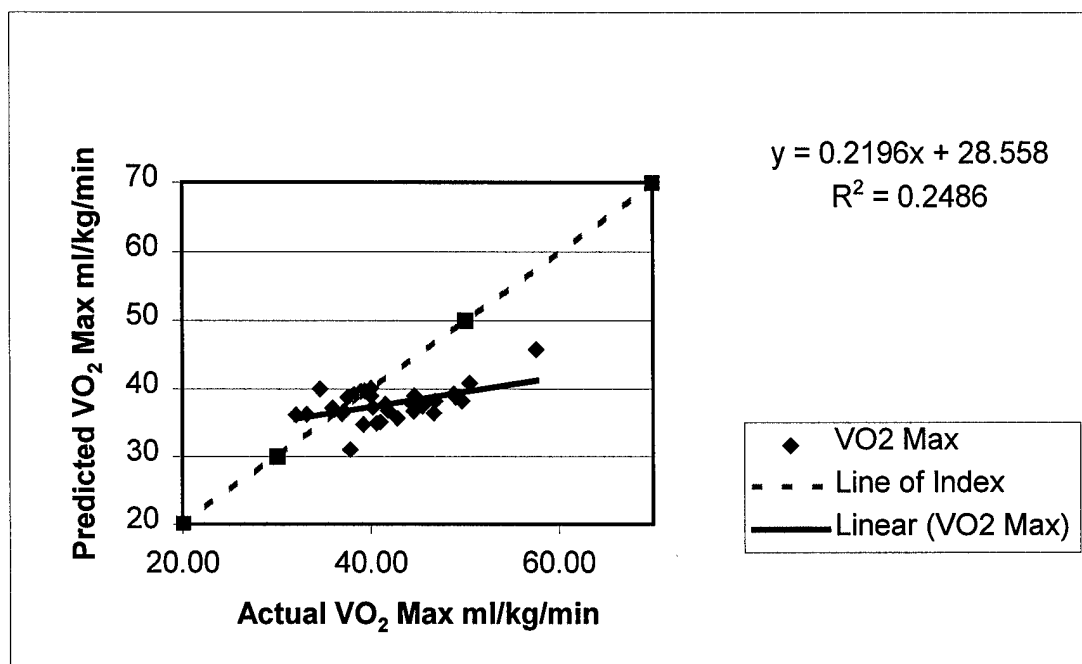


Figure 6: Actual  $\text{VO}_2 \text{ Max}$  vs. Predicted  $\text{VO}_2 \text{ Max}$  using Female Only Dolgener Equation ( $\text{L min}^{-1}$ )

This equation seems to predict the lower  $\text{VO}_2 \text{ max}$  of the population quite well, but the upper  $\text{VO}_2 \text{ max}$  values are under-estimated with this equation.

Another Dolgener equation evaluated was the general equation converted into the units of  $\text{ml/kg/min}$ . This equation is:

$$\text{VO}_2 \text{ max} = 88.7688 - (0.0957 \cdot \text{WT}) + (8.8924 \cdot \text{SEX}) - (1.4537 \cdot \text{T1}) - (0.1194 \cdot \text{HR})$$

SEX = 0:f, 1:m

WT = lbs

T1 = Time to complete walk (minutes)

HR = Final heart rate (bpm)

With this equation, a  $\text{VO}_2 \text{ max}$  mean of  $38.36 \text{ ml kg}^{-1} \text{min}^{-1}$  was predicted. A plot of these data against actual  $\text{VO}_2 \text{ max}$  (Figure 7) illustrates that for subjects at the lower end

of actual  $\text{VO}_2$  max the equation accurately predicted  $\text{VO}_2$  max, whereas the upper  $\text{VO}_2$  max values were under-estimated.

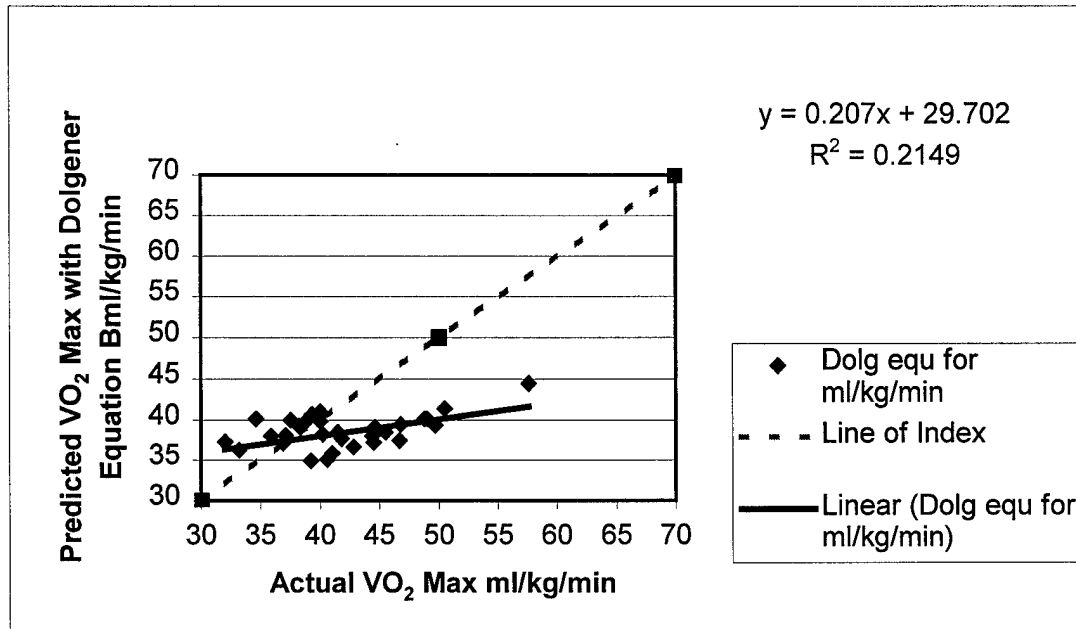


Figure 7: Predicted vs. Actual  $\text{VO}_2$  Max using Dolgener Equation for  $\text{ml kg}^{-1} \text{min}^{-1}$

The final applicable equation from the Dolgener study was the female-specific equation calculated in  $\text{ml kg}^{-1} \text{min}^{-1}$ . This equation is:

$$\text{VO}_2 \text{ max} = 79.4528 - (0.0743 \cdot \text{WT}) - (1.2027 \cdot \text{T1}) - (0.1011 \cdot \text{HR})$$

WT = lbs

T1 = Time to complete walk (minutes)

HR = Final heart rate (bpm)

This equation predicted a mean  $\text{VO}_2$  max of  $38.057 \text{ ml kg}^{-1} \text{min}^{-1}$ . A plot of these data (Figure 8) again illustrates that the lower  $\text{VO}_2$  max portion of the population is predicted well while the upper portion of the population is under-estimated.

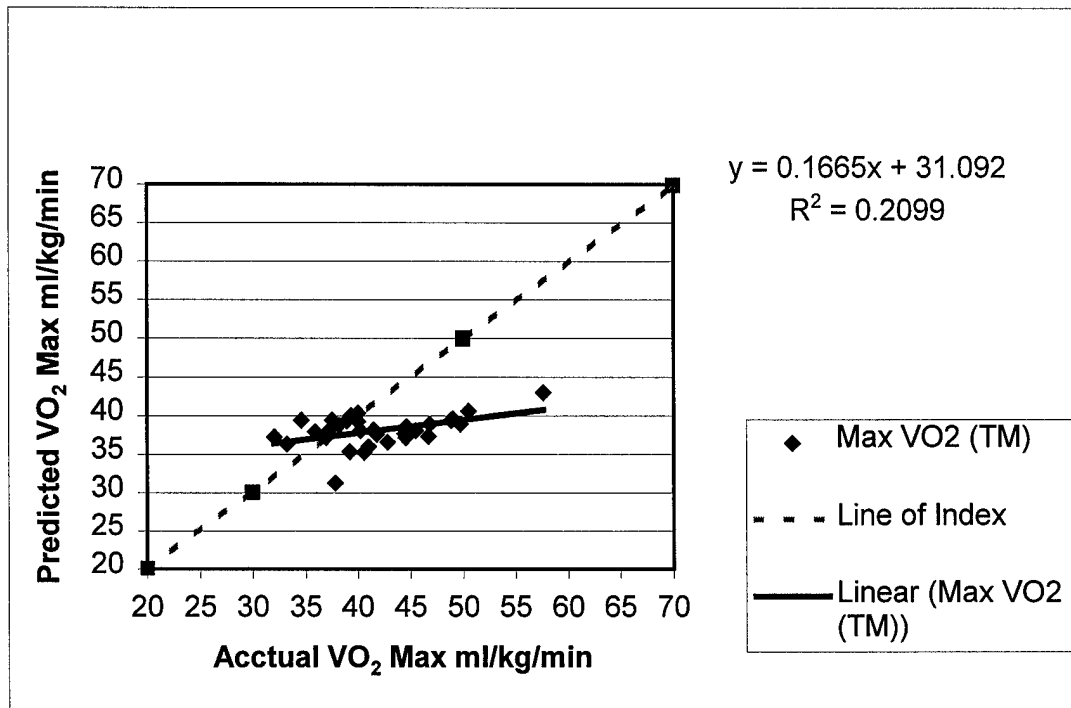


Figure 8: Actual  $\text{VO}_2$  Max vs. Predicted  $\text{VO}_2$  Max Using Dolgener Female Equation for  $\text{ml kg}^{-1} \text{min}^{-1}$

## SUMMARY

The Dolgener study provided 3 additional applicable equations that were evaluated in the present study once the general equation did not accurately predict  $\text{VO}_2$  max in the current sample of female military personnel. These additional equations were only reliable predictors of  $\text{VO}_2$  max in subjects with lower actual  $\text{VO}_2$  max values, but they underestimated  $\text{VO}_2$  max in subjects with a higher  $\text{VO}_2$  max.

## CHAPTER V

### DISCUSSION

The results of this study indicate that the 1-mile walk test, using the Dolgener et al. (1994) equation, is not a reliable predictor of  $\text{VO}_2$  max in a college-age female population in good physical condition. These results do not confirm findings of previous studies performed on this age group (Chuba 2000, George et al. 1998, Dolgener et al. 1994). Each of these earlier studies indicated that the Kline et al. (1987) equation over-predicts  $\text{VO}_2$  max in male and female college populations. Dolgener et al. (1994) reported over-predictions of 16-18% in males and 22-23% in females, while George et al. (1998) reported over-predictions of 8-20%, and Chuba (2000) found over-predictions ranging from 8-28% for a predominately male college-age population. These studies concluded that the equations by Dolgener et al. (1994) were better at predicting  $\text{VO}_2$  max in college-age populations.

Over-predictions using the Kline et al. (1987) equation were attributed by Dolgener et al. (1994) to the subjects' fitness levels. The youngest subject group in the original Kline et al. (1987) study (ages 30-39) had an average  $\text{VO}_2$  max of  $43 \text{ ml kg}^{-1} \text{ min}^{-1}$ , whereas the Dolgener et al. (1994) population (age 18-29) had an average  $\text{VO}_2$  max of  $41.2 \text{ ml kg}^{-1} \text{ min}^{-1}$ . An independent sample t-test indicates that this difference in  $\text{VO}_2$  max is significant ( $t(24) = -4.42$ ;  $p < 0.01$ ). George et al. (1998) stated that their subjects' aerobic capacity was  $42.8 \text{ ml kg}^{-1} \text{ min}^{-1}$ . The George study, which had a mean  $\text{VO}_2$  max closer to the Kline study, found similar results to the Dolgener study. In the



George study, it was concluded that the Kline equation over-predicted  $\text{VO}_2$  max.

Through an independent sample t-test, the subjects in the George et al. (1998) study were significantly different from the subjects in the original Kline study with values of  $t(129) = -0.98$  df;  $p < 0.05$ ). The subjects in the current study, although all women, had  $\text{VO}_2$  max values of  $41.83 \text{ ml kg}^{-1} \text{ min}^{-1}$ , which is similar to that observed in the Dolgener et al. (1994) study (as seen in Figure 3 and 4). However, Dolgener reported the females in his study to have a  $\text{VO}_2$  max of  $37.3 \text{ ml kg}^{-1} \text{ min}^{-1}$ . This is lower than the average of the current study and may be of consideration when using the Dolgener equation on the current population. This may be due to the Dolgener equation not being sensitive for higher  $\text{VO}_2$  max values, as seen in this study's population, and may be the reason for the over-estimations using the general Dolgener equation.

Twenty-seven of the 31 subjects (87%) in this study were USAFA cadets. Harger and Ellis (1975) illustrated that USAFA male cadets have "an excellent level of cardio-respiratory fitness especially when compared with other military samples within the same age span." At the time of their study, there were no female cadets enrolled at the USAFA, but since many of the current female subjects were intercollegiate athletes, one could speculate that these female cadets would, like their male counterparts evaluated in 1975, confirm the observations of Harger and Ellis. Harger and Ellis (1975) found that the male non-intercollegiate cadets ( $n=79$ ) had a higher aerobic capacity than that of the average college age population. In addition, when adjustments were made for residence at an altitude of 2,200 m, the  $\text{VO}_2$  values determined on a cycle ergometer increased from 47.91 to 51.75-55.09  $\text{ml kg}^{-1} \text{ min}^{-1}$ . These values would probably have been higher had the testing been conducted on a treadmill instead of a cycle ergometer. The average

actual  $\text{VO}_2$  max of  $41.8 \text{ ml kg}^{-1} \text{ min}^{-1}$  in the females in the current study was not adjusted for altitude. An altitude conversion would have increased the actual  $\text{VO}_2$  max by approximately 12%, since Tucker et al. (1984) found that with each 1000 m increment in altitude there is a 6% decrement in  $\text{VO}_2$  max. Furthermore, altitude exposure results in higher heart rates at the same absolute workload compared to sea level exercise. This would then produce an under-estimation of  $\text{VO}_2$  at altitude because the heart rates are elevated at altitude for any given workload. An altitude correction may bring the Dolgener predictions closer to the actual value. However, more research is needed to address a possible altitude conversion in order to compare studies conducted at moderate altitude with those conducted at sea level.

Harger and Ellis (1975) compared  $\text{VO}_2$  max of the cadets in their study to a study of USAF personnel of similar ages and found that  $\text{VO}_2$  max of non-cadet USAF personnel to be around  $40 \text{ ml kg}^{-1} \text{ min}^{-1}$  (Froelicher et al. 1974). This  $\text{VO}_2$  max value is significantly lower than the Harger and Ellis value ( $t(121) = 17.9 \text{ } p < 0.01$ ). However, this value is close to the value in the Dolgener study, which is also significantly lower than the Harter and Ellis values ( $t(351) = 19.3 \text{ } p < 0.01$ ). The current study's female population is comparable to that of the USAF personnel because of their closely related mean  $\text{VO}_2$  max values. Therefore, the females in the study were comparable to the  $\text{VO}_2$  max value of their corresponding active duty mixed-gender counterparts that they are supposed to represent. It can also be assumed that the Dolgener study has a comparable subject population to that of the USAF.

Physical performance is dependent on several factors and motivation, in particular, may play a role in the success of the prediction equation. Rintala et al. (1992)

suggested that motivation may be an important factor in the success of the 1-mile walk test. The USAFA is a competitive environment, and though the subjects in this study usually walked on the track alone, it was obvious that competing for better times was a factor among the cadets. Therefore, this self-motivation assisted the majority of the cadets in maintaining a maximal pace throughout the walk. Dolgener et al. (1994) indicated that, theoretically, the pace of the walk should not affect the prediction of  $\text{VO}_2$  max, as long as the pace was sufficient to keep heart rate within the linear portion of the heart rate- $\text{VO}_2$  relationship. Golding et al. (1989) stated that this relationship starts with a heart rate of 110 bpm, and subjects in better shape may not be able to attain a heart rate above 110 by simply walking. Thus, it was suggested that fit subjects should perform a higher intensity field test such as running. All the females in the current study achieved heart rates over 110 bpm. Because the subjects in this study were very fit women, Rintala et al. (1992) may be correct in that motivation is a key factor in the reliability of this test as a predictor of  $\text{VO}_2$  max.

After finding that the Dolgener equation for predicting  $\text{VO}_2$  max over-predicted performance, both in  $\text{L min}^{-1}$  as well as converted  $\text{ml kg}^{-1} \text{min}^{-1}$  units, alternative equations (Dolgener et al. 1994) were evaluated. The equations illustrated that the lower  $\text{VO}_2$  max end of the population was predicted quite well with the alternative equations but the upper  $\text{VO}_2$  max end, or more fit portion of the subject population, was under-predicted by the equations. Speculation as to why these alternative equations produced variable results may be due to the fact that the Dolgener study's mean  $\text{VO}_2$  max was  $36.6 \text{ ml kg}^{-1} \text{min}^{-1}$  for the validation group and  $37.3 \text{ ml kg}^{-1} \text{min}^{-1}$  for the cross-validation

group. Table 4 illustrates that there were differences in not only  $\text{VO}_2$  max between the females of the Dolgener study but in other areas.

TABLE 4. Descriptive data from the three studies (mean  $\pm$  SD)

	<u>Dolgener VG</u>	<u>Dolgener CVG</u>	<u>Present</u>
Number of Subjects*	100	45	31
Age (yrs)	19.4 $\pm$ 3.08	19.1 $\pm$ 2.80	21.0 $\pm$ 2.70
Weight (kg)	61.6 $\pm$ 8.72	62.9 $\pm$ 6.42	63.29 $\pm$ 8.27
$\text{VO}_2$ max (ml/kg/min)	36.6 $\pm$ 4.73	37.3 $\pm$ 5.86	41.83 $\pm$ 5.65
HR Walk (bpm)	161 $\pm$ 16.1	159 $\pm$ 18.9	148.13 $\pm$ 20.48
Time Walk (min)	13.78 $\pm$ 0.87	13.80 $\pm$ 1.21	13.27 $\pm$ 1.14

\* All subjects are Females

Both of the  $\text{VO}_2$  max values were below the current study's value, which was 41.83 ml  $\text{kg}^{-1} \text{min}^{-1}$ . The current study  $\text{VO}_2$  max was not adjusted for altitude, which if converted would have increased each value by approximately 12%. Therefore, the predicted  $\text{VO}_2$  max from the general equation still slightly over-predicts the fitness of the current subjects despite an altitude conversion. In addition the current's study had lower heart rate value and lower walk time values. This illustrates that the current study's females are able to walk faster and still have lower heart rates. This illustrates again that this study's females are even more fit than the population in Dolgener's study and that the Dolgener equation is not equipped to predict the extremes of the population. It is possible that the current study's female subjects, who had  $\text{VO}_2$  max near the mean of the Dolgener study, would have better predicted their  $\text{VO}_2$  max with the Dolgener equation and the altitude conversion than those in the upper portion of the population group.

## CHAPTER VI

### CONCLUSIONS AND RECOMMENDATIONS

The present study does not support previous research indicating that the 1-mile walk test is a valid predictor of  $\text{VO}_2$  max in the female population aged 18-30 years when using the equation derived by Dolgener et al. (1994). The results of this study indicate that the USAF may wish to employ caution if deciding to implement the 1-mile walk test in their Reserve and National Guard units because, although Chuba (2000) found that the Dolgener et al. (1994) equation works in a predominately male population, the general Dolgener equation significantly overestimated  $\text{VO}_2$  max in a comparable population in the present study.

This study may need to be performed on a less fit college age female population in order to determine if the equation will predict  $\text{VO}_2$  max more accurately in this female population. The present study's female population's  $\text{VO}_2$  max is similar to the Dolgener and USAF personnel's mixed-gender  $\text{VO}_2$  max mean values, but less than the average of male cadets at USAFA. This illustrates that these females without male counterparts to inflate their  $\text{VO}_2$  max mean scored near their mix-gender counterparts indicating a high fitness level of this female population. It would be relevant to developing USAF fitness standards to repeat the Harger and Ellis (1975) study separating the genders and comparing the  $\text{VO}_2$  max values of both female and male cadets to their representative civilian and non-Academy USAF counterparts. It would be of interest to compare a mixed gender sample of USAFA cadets to a respective college age population. The

results of these studies would help to validate that USAFA cadets are more aerobically fit than civilian and non-Academy personnel of similar age and gender. This then could be of use in comparing results in normal college age populations to a known healthy and fit population group.

Another area that needs to be researched is the altitude and its effects on the exercising USAFA cadets. With the high fitness level of this population, it might be helpful to demonstrate altitude and its effects on their performance. In addition, the cadets reside at 2,200 m but maintain a very active fitness routine and may provide a select population for research in high altitude and exercise studies.

## REFERENCES

1. American College of Sports Medicine. (1995). **ACSM's Guidelines for Exercise Testing and Prescription**, Baltimore, MD: Williams and Wilkins. 29-48.
2. Astrand, P. O., Ryhming, I. (1954). A Nomogram for Calculations of Aerobic Capacity (Physical Fitness) from Pulse Rate During Sub-maximal Work. **Journal of Applied Physiology**, 7, 218-221.
3. Cahalin, L., Pappagianopoulos, P., Prevost, S., Wain, J., and Ginns, L. (1995). The Relationship of the Six-Min Walk Test to Maximal Oxygen Consumption in Transplant Candidates with End-Stage Lung Disease. **Chest**, 108, 452-459.
4. Chin, D. L., & Blackwood, G. V. (1996). Ergometry as a Predictor of Basic Military Training Success. **Military Medicine**, 161, 2, 75-77.
5. Chuba, D. M. (2000) **Validation of the One-mile Walk as a Predictor of Aerobic Capacity**.
6. DeVries, H. A., Klafs, C. E. (1965). Prediction of Maximal Oxygen Intake from Sub-maximal Tests. **Journal of Sports Medicine and Physical Fitness**, 5, 207-214.
7. Dolgener, F. A., Hensley, L. D., Harsh, J. J., and Fjelstul, J. K. (1994). Validation of the Rockport Fitness Walking Test in College Males and Females. **Research Quarterly for Exercise and Sport**, 65,2, 152-158.
8. Doolittle, T. L., Bigbee, R. (1968). The Twelve-Minute Run-Walk: A Test of Cardiorespiratory Fitness of Adolescent Boys. **Research Quarterly**, 39, 491-495.
9. Enright, P. L., Sherrill, D. L. (1998). Reference Equations for the Six-Minute Walk in Healthy Adults. **American Journal of Respiratory Critical Care Medicine**, 158, 1384-1387.
10. Fenstermaker, K. L., Plowman, S. A. and Looney, M. A. (1992). Validation of the Rockport Fitness Walking Test in Females 65 Years and Older. **Research Quarterly for Exercise and Sport** 63,3, 322-327.

11. Fernall, B., Kohrt, W. (1990). The Effect of Training Specificity on Maximal and Submaximal Physiological Responses to Treadmill and Cycle Ergometry. **Journal of Sports Medicine and Physical Fitness, 30, 268-275.**
12. Fitchett, M. A. (1985). Predictability of VO<sub>2</sub> Max from Submaximal Cycle Ergometer and Bench Stepping Tests. **British Journal of Sports Medicine, 19, 85-87.**
13. Froelicher, V. F., Allen, M., Lancaster, M. C. (1974). Maximal Treadmill Testing of Normal USAF Aircrewman. **Aerospace Medicine, 45, 3, 310-315.**
14. George, J. D., Fellingham, G. W. and Fisher, A. G. (1998). A Modified Version of the Rockport Fitness Walking Test for College Men and Women. **Research Quarterly for Exercise and Sport 69,2, 205-209.**
15. Glassford, R. G., Baycroft, G. H. Y., Sedgwick, A. W., Macnab, R. B. J. (1965). Comparison of Maximal Oxygen Uptake Values Determined by Predicted and Actual Methods. **Journal of Applied Physiology 20,3, 509-513.**
16. Golding, L. A., Meyers, C. R., and Sinning, W. E. (1989). The Y's Way to Physical Fitness. Rosemont, IL: YMCA of America.
17. Harger, B. S., Ellis, R. P. (1975). Circulo-Respiratory Fitness in United States Air Force Academy Cadets. **Aviation, Space, and Environmental Medicine, 46,9 1144-1146.**
18. Harrison, M. H., Bruce, D. L., Brown, G. A., Cochrane, L. A. (1980). A Comparison of Some Indirect Methods for Predicting Maximal Oxygen Uptake. **Aviation, Space, and Environmental Medicine 51,10, 1128-1133.**
19. Harrison, M. H., Brown, G. A., Cochrane, L. A. (1980). Maximal Oxygen Uptake: Its Measurement, Application, and Limitations. **Aviation, Space Environmental Medicine 51,10, 1123-1127.**
20. Hartung, H. G., Krock, L. P., Crandall, C. G., Bisson, R. U., Myhre, L. G. (1993). Prediction of Maximal oxygen Uptake from Sub-maximal Exercise Testing in Aerobically Fit and Non-fit Men. **Aviation, Space and Environmental Medicine, 64, 735-740.**
21. Issekutz, B. Jr., Birkhead, N. C., Rodahl, K. (1962). Use of Respiratory Quotients in Assessment of Aerobic Work Capacity. **Journal of Applied Physiology, 17, 1, 47-50.**
22. Jackson, A., Weduwe, K. D., Schick, R., Sanchez, R. (1990). An Analysis of the Validity of the Three-mile Run as a Field Test of Aerobic Capacity in College Males. **Research Quarterly, 61,3, 233-237.**



23. Jessup, G. T., Tolson, H., Terry, J. W. (1974). Prediction of Maximal oxygen Intake from Astrand-Rhyming Test, 12-minute Run, and Anthropometric Variables Using Stepwise Multiple Regression. **American Journal of Physical Medicine, 53, 4, 200-207.**
24. Kline, G. M., Porcari, J. P., Hintermeister, R., Freedson, P. S., Ward, A., Mccarron, R. F., Ross, J., and Rippe, J. M. (1987). Estimation of VO<sub>2</sub> max from a One-Mile Track Walk, Gender, Age, and Body Weight. **Medicine and Science in Sports and Exercise, 19,3, 253-259.**
25. Kline, G. M., Porcari, J., Freedson, P., Ward, A., Ross, J., Wilkie, S., Rippe, J. (1987). Does Aerobic Capacity Affect the Validity of the One-Mile Walk VO<sub>2</sub> Max Prediction? **Medicine and Science in Sports and Exercise 19,2, S29, 172.**
26. Knapik, Joseph. (1989). The Army Physical Fitness Test (APFT): A review of the Literature. **Military Medicine, 154, 6, 326-329.**
27. Lockwood, Par A., Yoder, James E., & Deuster, Patricia A. (1997). Comparison and Cross-Validation of Cycle Ergometry Estimates of VO<sub>2</sub> max. **Medicine and Science in Sports and Exercise, 29,11, 1513-1520.**
28. Maksud, Michael G., Coutts, Kenneth, D. (1971). Application of the Cooper Twelve-Minute Run-Walk Test to Young Males. **Research Quarterly, 42, 54-59.**
29. McArdle, William, D., Katch, Frank, I., Katch, Victor, L. (1996). **Exercise Physiology**, Baltimore, MD: Williams and Wilkins. 126, 145-159,199-202,254-255,439,635.
30. Mitchell, Jere, H., Sproule, Brian, J., Chapman, Carleton, B. (1957). The Physiological Meaning of the Maximal Oxygen Intake Test. **Journal of Clinical Investigation, 37, 538-547.**
31. Myles, W. S., Toft, R. J., (1982). A Cycle Ergometer Test of Maximal Aerobic Power. **European Journal of Applied Physiology, 49, 121-129.**
32. Oja, P., Laukkanen, R., Pasanen, M., Tyry, T., Vuori, I. (1991). A 2-km Walking Test for Assessing the Caridiorespiratory Fitness of Healthy Adults. **International Journal of Sports Medicine, 12, 4, 356-362.**
33. Patton, J. F., Vogel, J. A., Mello, R. P. (1982). Evaluation of a Maximal Predictive Cycle Ergometer Test of Aerobic Power. **European Journal of Applied Physiology, 49, 131-140.**
34. Rintala, P., Dunn, J. M., Mccubbin, J. A., & Quinn, C. (1992). Validity of a Cardiorespiratory Fitness Test for Men with Mental Retardation. **Medicine and Science in Sports and Exercise, 24, 8, 941-945.**

35. Roul, Gerald, Germain, Philippe, & Bareiss, Pierre. (1998). Does the Six-Minute Walk Test Predict the Prognosis in Patients with NYHA Class II or III Chronic Heart Failure? **American Heart Journal, 136, 3, 449-457.**
36. Rowell, L. B., Taylor, H. L., Wang, Y. (1964). Limitations to Prediction of Maximal oxygen Intake. **Journal Applied Physiology, 19, 5, 919-927.**
37. Sharp, John R., (1991). The New Air Force Fitness Test: A Field Trial Assessing Effectiveness and Safety. **Military Medicine, 156, 4, 181-185.**
38. Teraslinna, P., Ismail. A. H., Macleod, D. F. (1996). Nomogram By Astrand and Ryhming as a Predictor of Maximal Oxygen Intake. **Journal of Applied Physiology, 21, 2, 513-515.**
39. Tucker, Alan, Stager, J. M., Cordain, L. (1984). Arterial O<sub>2</sub> Saturation and Maximum O<sub>2</sub> Consumption in Moderate-Altitude Runners Exposed to Sea Level and 3,050 m. **JAMA, 252, 20, 2867-2871.**
40. Williford, Henry N., Sport, Kim, Wang, Naizhen B. M., Olson, M. S., Blessing, D. (1994). The Prediction of Fitness Levels of United States Air Force Officers: Validation of Cycle Ergometry. **Military Medicine, 159, 3, 175-178.**
41. Zwiren, L. D., Freedson, P. S., Ward, A., Wilke, S., Rippe, J. M. (1991). Estimation of VO<sub>2</sub> Max: A Comparative Analysis of Five Exercise Tests. **Research Quarterly for Exercise and Sport, 62, 1, 73-78.**